

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-204250

(43)Date of publication of application : 19.07.2002

(51)Int.Cl.

H04L 12/56

(21)Application number : 2000-400634

(71)Applicant : FUJITSU LTD

(22)Date of filing : 28.12.2000

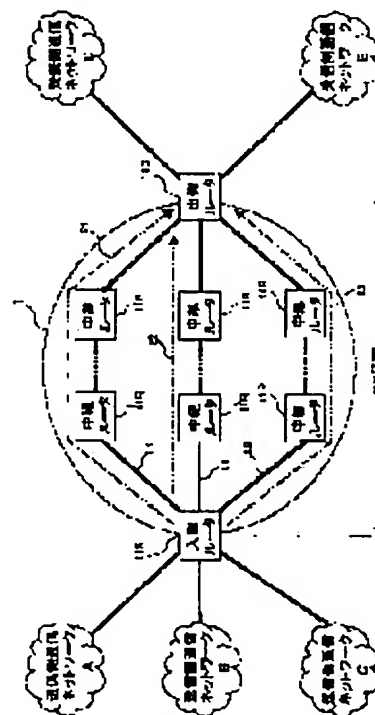
(72)Inventor : NAKAMICHI KOJI
TAKASHIMA KIYONARI
SOMIYA TOSHIO

(54) DEVICE AND METHOD FOR GATHERING TRAFFIC INFORMATION

(57)Abstract:

PROBLEM TO BE SOLVED: To enable a node on a communication network to gather information regarding traffic.

SOLUTION: The communication network 1 is a communication network where additional decentralization control by traffic engineering is performed. An in-side router 11S, a repeating router 11R, and an out-side router 11D arranged on the communication network 1 gather and store traffic information on respective links connected to themselves in a database. Those routers 11 floods the gathered traffic information by opaque LSA. Each router 11 having received the opaque LSA stores information on the opaque LSA in its database. Consequently, the in-side router 118 performs load decentralization control according to information on the traffic of respective links.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-204250

(P 2 0 0 2 - 2 0 4 2 5 0 A)

(43) 公開日 平成14年7月19日 (2002.7.19)

(51) Int.Cl. ⁷

H04L 12/56

識別記号

F I

H04L 11/20

102

D 5K030

テーマコード (参考)

審査請求 未請求 請求項の数 5 O L (全22頁)

(21) 出願番号 特願2000-400634 (P 2000-400634)

(22) 出願日 平成12年12月28日 (2000.12.28)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 仲道 耕二

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72) 発明者 高島 研也

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74) 代理人 100094514

弁理士 林 恒徳 (外1名)

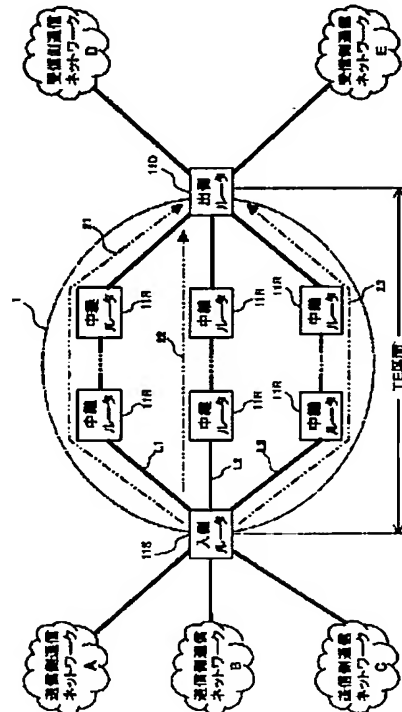
最終頁に続く

(54) 【発明の名称】 トラフィック情報収集装置およびトラフィック情報収集方法

(57) 【要約】

【課題】 通信ネットワーク上のノードがトラフィックに関する情報を収集できるようにする。

【解決手段】 通信ネットワーク1は、トラフィック・エンジニアリングによる付加分散制御が行われる通信ネットワークである。通信ネットワーク1に配置された入側ルータ11S、中継ルータ11Rおよび出側ルータ11Dは、自己に接続された各リンクのトラフィック情報を収集し、データベースに記憶する。また、これらの各ルータ11は、収集したトラフィック情報をオペークL SAによりフラッディングする。オペークL SAを受信した各ルータ11は、自己のデータベースにオペークL SAの情報を記憶する。これにより、入側ルータ11Sは、各リンクのトラフィックの情報に基づいて負荷分散制御を行う。



【特許請求の範囲】

【請求項 1】 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードに設けられ、前記通信路のトラフィックに関する情報を収集するトラフィック情報収集装置であって、自己のノードに接続された通信路のトラフィックに関する情報を収集するトラフィック情報収集部と、前記通信ネットワーク上の通信プロトコルで使用されるメッセージにより、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報を他のノードに送信するトラフィック情報送信部と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報を受信するトラフィック情報受信部と、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、前記トラフィック情報受信部により受信された前記他のノードに接続された通信路のトラフィックに関する情報とを記憶するトラフィック情報記憶部と、を備えているトラフィック情報収集装置。

【請求項 2】 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードに設けられ、前記通信路のトラフィック情報を収集するトラフィック情報収集装置であって、自己のノードに接続された通信路のトラフィックに関する情報を収集するトラフィック情報収集部と、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報とを記憶するトラフィック情報記憶部と、を備えているトラフィック情報収集装置。

【請求項 3】 請求項 1 または 2 において、前記他のノードから送信されてきたトラフィックに関する情報を、送信されてきた通信路を除く他の通信路に転送するトラフィック情報転送部をさらに備えている、トラフィック情報収集装置。

【請求項 4】 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードで行われる、前記通信路のトラフィック情報を収集するトラフィック情報収集方法であって、自己のノードに接続された通信路のトラフィックに関する情報を収集し、前記収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報とを記憶する、トラフィック情報収集方法。

【請求項 5】 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードで実行

される、前記通信路のトラフィック情報を収集するトラフィック情報収集プログラムであって、自己のノードに接続された通信路のトラフィックに関する情報を収集する手順と、前記収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報とを記憶する手順と、を備えているトラフィック情報収集プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、通信ネットワークにおけるトラフィック情報収集装置、トラフィック情報収集方法、およびトラフィック情報収集プログラムに関し、特に、通信ネットワークにおいて、負荷分散を行うために必要な情報を収集するためのトラフィック情報収集装置、トラフィック情報収集方法、およびトラフィック情報収集プログラムに関する。

【0002】また、本発明は、通信ネットワークにおいて、負荷分散を行うために必要な情報を収集する機能を有する、通信ネットワークにおけるノードに関する。

【0003】

【従来の技術】インターネットは、コネクションを確立しないコネクションレスなネットワークであり、通信データは、IP (Internet Protocol) アドレスが付与された IP パケットに分割され送信される。IP パケットは、IP アドレスに基づいてルータ間を中継され、宛先ルータまたはコンピュータ（以下、単に「宛先」という。）に届けられる。

【0004】IP アドレスに基づいてルータ間をどのように中継して、IP パケットを宛先まで届けるかは、ルーティング・プロトコルによって決定される。ルーティング・プロトコルとして、現在、RIP (Routing Information Protocol) または OSPF (Open Shortest Path First) が一般に使用されている。これらのルーティング・プロトコルは、基本的には、通信ネットワーク内の各リンクのコスト（たとえばホップ数）に基づいて宛先までの最短の経路（ルート、パス）を求め、その最短経路を IP パケットの通信経路として設定するものである。

【0005】

【発明が解決しようとする課題】しかし、これらのルーティング・プロトコルでは、コストに基づく最短経路が求められるだけであって、各リンクのトラフィックの状況を考慮した経路は求められない。しかも、経路が新たに設定されるのは、通信ネットワークのトポロジーに変化が生じた場合だけであり、リアルタイムに設定/変更が行われるわけではない。一方、インターネットの普及により、インターネット内を行き交うパケットの量（すなわちトラフィック量または負荷）は急激に増加している。

【0006】その結果、ルーティング・プロトコルにより設定された経路上で、輻輳が発生するようになっていく。

【0007】なるほど、OSPFには、複数の経路を設定してIPパケットを分散する等コスト・マルチパス(Equal Cost Multipath)というルーティングもある。しかし、このルーティングは、コストが等しい経路が複数存在する場合に限られており、また、OSPFを使用しているため、リアルタイムの通信ネットワークの負荷の変動に対処できない。

【0008】そこで、現在、インターネットにおける輻輳を回避する技術体系として、トラフィック・エンジニアリングが検討されている。このトラフィック・エンジニアリングは、一般に、宛先までの経路を複数設定するとともに、各経路のトラフィックの状況をリアルタイムに監視して、空きのある経路や負荷の小さな経路を選択してIPパケットを送信し、複数の経路間で負荷(トラフィック)の分散を行うものである。

【0009】一方、このようなトラフィック・エンジニアリングを行うためには、各ルータが、各経路の使用状況を含む、トラフィックに関する情報を知る必要がある。

【0010】本発明は、このような状況に鑑みなされたものであり、トラフィックの状況に応じて負荷分散を図ることができるように、通信ネットワーク上のノードがトラフィックに関する情報を収集できるようにすることにある。

【0011】

【課題を解決するための手段】前記目的を達成するために、本発明に係る情報収集装置は、通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードに設けられ、前記通信路のトラフィックに関する情報を収集するトラフィック情報収集装置であって、自己のノードに接続された通信路のトラフィックに関する情報を収集するトラフィック情報収集部と、前記通信ネットワーク上の通信プロトコルで使用されるメッセージにより、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報を他のノードに送信するトラフィック情報送信部と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報を受信するトラフィック情報受信部と、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、前記トラフィック情報受信部により受信された前記他のノードに接続された通信路のトラフィックに関する情報とを記憶するトラフィック情報記憶部と、を備えている。

【0012】また、本発明に係る情報収集装置は、通信される情報を送信、受信または中継する複数のノード

と、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードに設けられ、前記通信路のトラフィック情報を収集するトラフィック情報収集装置であって、自己のノードに接続された通信路のトラフィックに関する情報を収集するトラフィック情報収集部と、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報とを記憶するトラフィック情報記憶部と、を備えている。

【0013】本発明に係る情報収集装置によると、通信ネットワークにおけるノードは、自己のノードに接続された通信路のトラフィックに関する情報を収集するとともに、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報を受信し、これらの情報を記憶する。したがって、ノードは、記憶されたこれらの情報に基づいて、通信ネットワーク上の各通信路のトラフィックの状況を知ることができ、その結果、トラフィックの状況に応じた負荷分散制御を行うことが可能となる。

【0014】好ましくは、前記他のノードから送信されてきたトラフィックに関する情報を、送信されてきた通信路を除く他の通信路に転送するトラフィック情報転送部が、さらに備えられる。

【0015】これにより、あるノードに送信されてきた、他のノードのトラフィックに関する情報を、それ以外の他のノードにも伝達することができ、伝達された他のノードも、トラフィックに関する情報を入手することができる。その結果、通信ネットワーク上のすべてのノードが、該通信ネットワークの各通信路のトラフィックの状況を把握することができる。

【0016】

【発明の実施の形態】<通信ネットワーク1の構成>図1は、本発明に係る「情報収集装置」を含んだノードを有する通信ネットワーク1の概略構成を示すブロック図である。この通信ネットワーク1は、本実施の形態ではインターネットの一部を形成するとともに、トラフィック・エンジニアリング(以下「TE」という。)による負荷分散が行われる通信ネットワークである。また、この通信ネットワーク1は、本実施の形態では、非同期転送モード(Asynchronous Transfer Mode: 以下「ATM」という。)により、IPパケット(以下、単に「パケット」という。)を送受信する通信ネットワークである。

【0017】通信ネットワーク1には、外部の通信ネットワークとして、通信ネットワーク1にパケットを送信する送信側通信ネットワークA、BおよびCと、通信ネットワーク1から送信されるパケットを受信する受信側通信ネットワークDおよびEとが接続されている。

【0018】これら外部の通信ネットワークを「送信側

通信ネットワーク」と「受信側通信ネットワーク」とに分けたのは、TEを説明するための便宜上のためであり、送信側通信ネットワークはパケットの送信のみを行うとは限らず、また、受信側通信ネットワークはパケットの受信のみを行うとは限らない。たとえば、通信ネットワーク1から送信側通信ネットワークA、BおよびCにパケットが送信される場合もあるし、また、受信側通信ネットワークDおよびEから通信ネットワーク1にパケットが送信される場合もある。

【0019】これらの送信側通信ネットワークA、BおよびC、ならびに受信側通信ネットワークDおよびEは、インターネットの一部を形成し、本実施の形態では、イーサネット（登録商標）で構築されている。

【0020】通信ネットワーク1は、「ノード」の一例としてのルータ（後述するラベル・スイッチ・ルータ）を複数備えている。これらのルータには、入側ルータ11Sと、出側ルータ11Dと、複数の中継ルータ11Rとが含まれている。

【0021】入側ルータ11Sは、送信側通信ネットワークA、BまたはC（または入側ルータ11に接続された図示しないコンピュータ端末）からパケットを受信すると、そのパケットに対してTEを実行し、TEに基づいて設定された経路（パス、ルート）上の隣接した中継ルータ11Rに送信するものである。中継ルータ11Rは、入側ルータ11Sと出側ルータ11Dとの間に配置され、入側ルータ11Sからのパケットを中継して出側ルータ11Dに渡すものである。出側ルータ11Dは、入側ルータ11Sから中継ルータ11Rを介して送信されてきたパケットを受信側通信ネットワークDまたはE（または出側ルータ11Dに接続された図示しないコンピュータ端末）に送信するものである。

【0022】なお、ここでも、ルータを入側ルータ11Sと、出側ルータ11Dと、中継ルータ11Rとに分けたのはTEを説明するための便宜上のためであり、たとえば、入側ルータ11Sが出側ルータとなる場合もあるし、出側ルータ11Dが入側ルータとなる場合もある。また、以下では、これらの入側ルータ11S、出側ルータ11D、および中継ルータ11Rを区別しない場合には、単に「ルータ11」と総称することとする。

【0023】ルータ11間には、各ルータを接続する物理的な「通信路」（ないしは「伝送路」）としてのリンクが設けられている。たとえば、入側ルータ11Sには、隣接する3つの中継ルータ11Rとの間に、3つのリンクL1～L3がそれぞれ接続されている。

【0024】このように、入側ルータ11Sは、外部の送信側通信ネットワークA、BおよびC（ならびに入側ルータ11Sに接続されたコンピュータ端末）から送信されてきたパケットに対してTEを実行する始点となり、出側ルータ11Dは、TEの終点となる。したがって、入側ルータ11Sから出側ルータ11Dまでの区間

は、「TE区間」と呼ばれることがある。

【0025】通信ネットワーク1は、本実施の形態では、ラベル・スイッチ技術の一例としてのMPLS (Multi-Protocol Label Switching) を用いて経路を設定するように構成されている。このため、入側ルータ11S、出側ルータ11D、および中継ルータ11Rは、ラベル・スイッチ・ルータ (Label Switch Router: 以下「LSR」という。) として構成されている。

【0026】MPLSは、OSI参照モデルにおけるレイヤ3のインターネット・プロトコルのルーティング処理と、ATMやフレーム・リレー等のレイヤ2のスイッチング処理とを融合させる技術である。LSRは、パケット転送（パケット交換）情報として、IPアドレスよりも下位のレイヤの情報である「ラベル」を使用する。ATMスイッチをLSRのパケット転送エンジン（ラベル・スイッチ・エンジン）に使用する場合には、ラベルとしてVPI/VCI (Virtual Path Identifier/Virtual Connection Identifier) が使用される。すなわち、IPアドレスがラベル（VPI/VCI）にマッピングされ、ラベルを用いてパケット転送が行われる。

【0027】MPLSでは、LSP (Label Switched Path) と呼ばれるあらかじめ設定された経路（仮想コネクション）上でパケットを転送する。このLSPは、OSPFのような既存のルーティング・プロトコルにより定められた経路に沿って設定することもできるし、既存のルーティング・プロトコルによる経路とは独立に設定することもできる。また、宛先（ルータまたはコンピュータ端末）まで、複数のLSPを設定することもできる。

【0028】したがって、複数のLSPを既存のルーティング・プロトコルとは独立に設定し、これらのLSP間で負荷（トラフィック）の分散を行うことにより、TEが可能となる。

【0029】また、MPLSでは、最初から複数のLSPを設定しておくのではなく、通信ネットワーク1内のトラフィック量の状態を随時監視しておき、ある経路のトラフィック量が多くなった場合に、同じ宛先に向かう別の経路を検索して新たなLSPを設定し、このLSPにトラフィックを分散する制御も可能である。これによっても、通信ネットワークのトラフィックの状態に応じたダイナミックなTEが可能となる。

【0030】LSPは、1本のリンク内に複数個設定することができる。たとえば、100 [Mbps (Mビット/秒)] の帯域を有する1本のリンク内には、それぞれ20 [Mbps] のLSPを5個設定することができる。

【0031】LSPの設定には、たとえば、RSVP (Resource Reservation Protocol) またはこれを拡張したRSVP-tunnelを用いて、経路上のリンクの帯域を予約することにより行うことができる。

【0032】通信ネットワーク1では、このLSPによる経路（ラベル・スイッチ・パス、ATMにおける仮想コネクション）として、図1に示すように、3つの経路21～23があらかじめ設定され、または、動的に生成されるものとする。すなわち、入側ルータ11Sは、パケットを送信するための3つの経路21～23を有する。

【0033】各経路21～23は、1または2以上のリンク内に設けられたLSPから構成される。また、各経路21～23を構成するリンクは、すべて異なるリンクとは限らず、その一部または全部は同一リンク内に設けられた、異なるLSPにより構成される場合もある。

【0034】入側ルータ11Sは、TE区間外にある送信側通信ネットワークA、BおよびCとの境界に位置し、これらの送信側通信ネットワークA、BまたはCから送信されてきたパケットにラベルを付加し、中継ルータ11Rに送信する。中継ルータ11Rは、ラベルが付加されたパケットをラベル交換して、出側ルータ11Dに向けて転送する。したがって、中継ルータ11Rは、パケットのIPアドレスのルーティング・エントリを検索することなく、ラベル情報だけでパケットの転送（ラベル・スイッチ転送）を行うことができる。出側ルータ11Dは、TE区間外にある受信側通信ネットワークDおよびEとの境界に位置し、送信されてきたパケットからラベルを取り除き、受信側通信ネットワークDまたはE（または出側ルータ11Dのコンピュータ端末）にこのパケットを送信する。なお、通信ネットワーク1は、前述したように、ATMを採用しているため、ここでいうラベルはVPI/VCIとなり、またパケットはATMセルとして転送されることとなる。

【0035】通信ネットワーク1のトラフィックの状態に応じて、LSPによる経路を設定し、トラフィックの負荷分散を行うために、各ルータ11は、自己に接続されている各リンクおよび各リンク内の各LSPのトラフィックに関する情報（以下「トラフィックに関する情報」を単に「トラフィック情報」という。）を交換する。このトラフィック情報の交換には、本実施の形態では、OSPF（Open Shortest Path First）における特殊なLSA（リンク状態広告：Link State Advertisement）であるオpaque LSA（Opaque Link State Advertisement）が使用される。

【0036】オpaque LSAは、その中に含むことができる情報については、後述するLSAヘッダおよびLSRインタフェース情報を除いて特に規定がなく、自由に使用できる領域を有する。また、オpaque LSAは、その伝播（送信）手法としてフラッディング（Flooding）が使用される点を除いて、いつ伝播するかについても規定がなく、各ルータが所望の時に伝播することができる。しかも、オpaque LSAは、OSPFというプロトコルに従っているため、インターネット等の通信ネット

ワークにおいて使用することができる。したがって、自由に使用できる領域にトラフィック情報を含めて伝播することにより、インターネット等の既存の通信ネットワーク上の各ルータ間でトラフィック情報を交換することができる。

【0037】＜ルータ11の構成＞図2は、ルータ11を代表して、入側ルータ11Sの構成を示すブロック図である。入側ルータ11Sは、処理装置30と、記憶装置32と、インタフェース装置41～46とを備えている。また、入側ルータ11Sにコンピュータ端末が接続されている場合には、このコンピュータ端末用のインタフェース装置47が設けられる。なお、中継ルータ11Rおよび出側ルータ11Dは、インタフェース装置の個数が異なる場合があることと、その接続先が異なる場合があることとを除いて、入側ルータ11Sと同様の構成を有するので、ここでは、入側ルータ11Sの構成についてのみ説明することとする。

【0038】記憶装置32は、ハードディスク装置等で構成され、リンク状態データベース32aおよびオpaque LSA用データベース32bを含んでいる。リンク状態データベース32aは、OSPFで規定されている、リンク状態が格納されるデータベースである。オpaque LSA用データベース32bは、オpaque LSAが格納されるデータベースであり、その詳細な説明は後述する。

【0039】処理装置30は、図示しないCPU、メモリ（RAMおよびROM）等を備えている。この処理装置30は、その内部のメモリまたは記憶装置32に記憶されたプログラムに従って、記憶装置32およびインタフェース装置41～46（47）の制御を行うとともに、経路の設定、トラフィックの状態に応じた負荷分散制御、パケット（ATMセル）の送受信、オpaque LSAの作成および送受信、受信したLSAによるリンク状態データベースの更新、受信したオpaque LSAによるオpaque LSA用データベースの更新等の処理を行う。

【0040】インタフェース装置41～46（47）は、入出力バッファを備え、処理装置30の制御の下、接続された各リンクを転送されるパケット（ATMセル）の入出力処理を行う。

【0041】なお、パケット（ATMセル）の送受信（ATMセルの生成、ラベル（VPI/VCI）のスイッチングを含む。）は、ハードウェア回路により構成されたハードウェア交換機をルータ11S内に設けて、この交換機により行うことができる。

【0042】＜オpaque LSAのデータ構造＞トラフィック情報は、前述したように、OSPFにおけるオpaque LSAを用いてルータ11間をフラッディングされる。このオpaque LSAは、図3に示すデータ構造を有する。図中、最上部に示す数字はビット番号である。したがって、この図3では、オpaque LSAを横方向に3

2ビット（4バイト）単位で整列して示している。

【0043】オペークLSAは、LSAヘッダおよびLSRインタフェース情報のフィールドを除いては、特に規定がなく、自由に使用できるフィールド（図3のリンク情報のフィールド）を有する。したがって、本実施の形態では、図3に示すデータ構造により、トラフィック情報がルータ11間をフラディングされる。

【0044】オペークLSAは、大きく分けると、LSAヘッダとオペーク・フィールドとに分けられる。

【0045】LSAヘッダは、20バイトからなり、通常のLSA（オペークLSA以外のLSA）と同様のデータ構造を有する。すなわち、LSAヘッダは、リンク状態エージ（Link State Age）、オプション（Option）、リンク状態タイプ（Link State Type: LSType）、オペーク・タイプ（Opaque Type）、オペーク識別子（Opaque Id）、広告ルータ（Advertising Router）、リンク状態番号（LS Sequence Number）、リンク状態チェック・サム（LS Checksum）、および長さ（Length）の各フィールドから構成されている。

【0046】「リンク状態エージ」は、2バイトからなり、このオペークLSAの寿命（秒単位）、すなわち有効期限を示している。

【0047】「オプション」は、1バイトからなり、ルータがオプション機能をサポートするように設定したり、そのサポート・レベル他のルータに伝えることを可能にするフィールドである。この「オプション」フィールドの第2ビットは、0ビットと呼ばれ、そのルータがオペークLSAをサポートしている（Opaque-capable）かどうかを示している。

【0048】「リンク状態タイプ」は、1バイトからなり、LSAのタイプを示している。オペークLSAの場合に、リンク状態タイプの値は、9、10または11となり、この値によってフラディングされる範囲が異なる。

【0049】この値が9の場合には、そのオペークLSAのフラディング範囲（FloodingScope）は“Link-local”であり、ローカル・ネットワーク内とされる。10の場合には、フラディング範囲は“Area-local”であり、そのオペークLSAは、接続しているエリアのボーダ（境界）を超えてフラディングされることはない。11の場合には、フラディング範囲は“Equivalent to AS-external LSA”であり、そのオペークLSAは、AS（Autonomous System）全体に亘ってフラディングされる。特に（1）すべての通過領域（Transit area）にフラディングされ、（2）バックボーン（backbone）からスタブ領域（stub area）にフラディングされず、（3）あるルータから、そのルータに接続されているスタブ領域に向けては生成されない。

$$\text{Ave_Utilization} = \alpha \times \text{CUTY}(n) + (1 - \alpha) \times \text{CUTY}(n-1) \quad \dots (1)$$

この式（1）は、移動平均（moving average）法による

【0050】「広告ルータ」は、4バイトとからなり、このオペークLSAを作成し、広告したルータのIPアドレスが置かれるフィールドである。リンク状態番号およびリンク状態チェック・サムは、同一の広告ルータが作成した2以上のオペークLSAがある場合に、いずれが最新（時間的に最も後）のものであるかを判定する際に用いられる。「長さ」は、2バイトとからなり、このオペークLSAの長さ（バイト数）を示している。

【0051】「オペーク・フィールド」は、LSRインタフェース情報と、そのルータ11に接続されたリンクの個数分設けられたリンク情報とから構成されている。

【0052】LSRインタフェース情報は、4バイトとからなり、上位2バイトのE_Bフィールドと、下位2バイトのリンク・カウン트의フィールドとから構成されている。「E_Bフィールド」は、そのルータがエリア・ボーダ・ルータかASバウンダリ・ルータかを示す。

「リンク・カウンタ」は、そのルータに接続されているリンクの個数を示す。

【0053】「リンク情報」は、リンク識別子（Link Id）、ネット・マスク、コネクション・タイプ、LSPカウンタ、およびリンク・データの各フィールドから構成されている。

【0054】「リンク識別子」は、4バイトとからなり、そのリンクが接続されている隣接ルータまたはコンピュータ端末のIPアドレスを示す。「ネット・マスク」は、4バイトとからなり、リンク識別子の（サブ）ネット・マスクを示す。「コネクション・タイプ」は、1バイトとからなり、そのリンクの接続先がルータか、コンピュータ端末か等を示す。LSPカウンタは、1バイトとからなり、そのリンクに設けられたLSPの個数を示す。この値は、通信時に設けられるLSPの個数が変化することに伴い動的に変化する。

【0055】「リンク・データ」は、複数のデータから構成されている。図4は、リンク・データのデータ構造を示している。リンク・データは、そのリンクのリンク統計情報と、そのリンク上に形成された各LSPのLSP統計情報とを有する。これらのリンク統計情報およびLSP統計情報は、本発明に係る「トラフィックに関する情報（トラフィック情報）」の一例である。

【0056】「リンク統計情報」は、平均使用率（Ave_Utilization）と、出力リンク廃棄パケット数（Loss）と、出力リンク帯域（BwI）との各フィールドを有する。

【0057】「平均使用率」は、4バイトとからなり、そのリンクの平均使用率を示す。この平均使用率は、以下の式により計算される。

【0058】

計算式である。ここで、 α は平滑化（スムージング）係

数である。CUTY (n) は現在 (すなわち時刻 n) のリンク使用率を、CUTY (n-1) は 1 時刻前 (すなわち時刻 (n-1)) のリンク使用率を、それぞれ示し、以下の

$$\text{CUTY (n)} = \frac{\text{時刻 (n-1) から時刻 n までのリンクへの出力パケット数}}{\text{出力リンク帯域}} \quad \dots (2)$$

なお、平均使用率は、観測された CUTY の最大値とするこ
ともでき、この場合に、平均使用率は以下の式で表され
る。

【0060】

$$\text{Ave_Utilization} = \max (\text{観測値}) \quad \dots (3)$$

「出力リンク廃棄パケット数」は、4 バイトからなり、
そのリンクの廃棄された出力パケットの合計数を示す。
出力リンク廃棄パケット数は、以下の式により計算され
る。

【0061】

$$\text{Loss (n)} = \text{Loss (n-1)} + \text{NLoss} \quad \dots (4)$$

ここで、Loss (n-1) は、1 時刻前の時刻 (n-1)
までの出力リンク廃棄パケットの合計数であり、NLoss
は、時刻 (n-1) から時刻 n までに生じた出力リンク
廃棄パケット数である。

【0062】「出力リンク帯域」は、4 バイトからな
り、出力リンクの帯域 [bps (ビット/秒)] を示
す。この出力リンク帯域は、各ルータ 11 の処理装置 3
0 のメモリまたは記憶装置 32 にあらかじめ記憶され、
記憶された値がオペーク LSA の該フィールドに書き込
まれるようになっている。

【0063】「LSP 統計情報」は、LSP 識別子 (ls
p_id) と、LSP_COS クラス (Cos_lsp) と、LSP
平均使用率 (Ave_Utility_lsp) と、LSP 出力廃棄
パケット数 (Loss_lsp) と、LSP 帯域 (Bw_lsp) と
の各フィールドを有する。

【0064】「LSP 識別子」は、2 バイトからなり、
その LSP の識別子を示す。「LSP_COS クラス」
は、2 バイトからなり、その LSP の COS クラスを示
している。

【0065】「LSP 平均使用率」は、4 バイトからな
り、その LSP の平均使用率を示し、前記式 (1) およ
び (2) において、出力リンクを、リンクではなく LSP
に置換した式により求められる。また、前記式 (3)
と同様に、その LSP における観測値の最大値とするこ
ともできる。

【0066】「LSP 出力廃棄パケット数」は、4 バイ
トからなり、その LSP の廃棄された出力パケットの合
計数を示し、前記式 (4) において、出力リンクを LSP
に置換することにより求められる。

【0067】「LSP 帯域」は、その LSP が設定され
る際に割り当てられた帯域 [bps] を示している。こ
の LSP 帯域は、LSP の設定時に、その LSP が設定
されるリンクを出力リンクとするルータ 11 の処理装置
30 のメモリまたは記憶装置 32 に記憶され、記憶され

式により求められる。

【0059】

た値がオペーク LSA の該フィールドに書き込まれるよ
うになっている。

【0068】これまでに述べたリンク平均使用率、出力
リンク廃棄パケット数、LSP 平均使用率、および LSP
出力廃棄パケット数は、たとえば、各ルータ 11 のオ
ペレーティング・システム (OS: たとえば Linu
x) がリアルタイムで収集し、管理しているシステム状
態 (CPU 負荷、転送パケット数等) を表すデータ (フ
ァイル) に基づいて求められる。

【0069】図 5 および図 6 は、ルータ 11 S の OS が
リアルタイムで収集/管理しているシステム状態のデー
タの一例を示している。図 5 はルータ 11 S に設けられ
た記憶装置 32 のディレクトリ "/proc/net/dev" にあ
るファイルの内容を、図 6 は記憶装置 32 のディレク
トリー "/proc/atm/device" にあるファイルの内容を、そ
れぞれテーブル (表) 形式により示している。

【0070】これらの図に示す数値は、ある時刻のデー
タの一例である。また、これらの数値は、ルータ 11 S
に電源を投入してからの積算値を示し、OS がリアルタ
イム (一定時間間隔 Δt) で収集しているので、時々刻
々と変化するものである。これらの数値は積算値である
ので、たとえば時刻 (n-1) にこのテーブルを参照し
た時の値と時刻 n にこのテーブルを参照した時の値との
差分を求めることにより、時刻 (n-1) から時刻 n ま
での 1 時刻の間の変化量を求めることができる。

【0071】図 5 に示すテーブルの上半分は、"Receiv
e" の文字が示すようにルータ 11 S の受信に関するデー
タを示し、下半分は、"Transmit" の文字が示すよう
にルータ 11 S の送信に関するデータを示す。以下、テ
ーブルの上半分を「受信テーブル」といい、下半分を
「送信テーブル」という。

【0072】受信テーブルおよび送信テーブルにおける
"interface" は、ルータ 11 S に接続されているイン
タフェースの種類を示している。このインタフェースに
は、"lo", "eth0", "eth1", "eth2", "atm
0", "atm1", および "atm2" がある。

【0073】"lo" は、"loopback device" の略であ
り、ルータ 11 S に接続されたコンピュータ端末 (図 1
には図示略) とのインタフェースを意味する。"eth0"
~ "eth2" は、イーサネットとのインタフェースを意味
する。ルータ 11 S には、図 1 に示すように、送信側通
信ネットワーク (イーサネット) A ~ C が接続されてい
るので、"eth0" は送信側通信ネットワーク A との、
"eth1" は送信側通信ネットワーク B との、"eth2" は
送信側通信ネットワーク C との、各インタフェースを意

味する。

【0074】“atm0”～“atm2”は、ATMとのインタフェースを意味する。図1に示すように、入側ルータ11Sには、ATMネットワークで構成された通信ネットワーク1の3つのリンクL1～L3が接続されているので、“atm0”はリンクL1との、“atm1”はリンクL2との、“atm2”はリンクL3との各インタフェースを意味する。

【0075】受信テーブルおよび送信テーブルの“bytes”は、それぞれ、各インタフェース(interface)から受信したバイト数および各インタフェースへ送信したバイト数を示している。たとえば、インタフェース“lo”のデータによると、入側ルータ11Sは、コンピュータ端末から6084バイトのデータを受信し、同じバイト数のデータをインタフェース“lo”に送信していることが示されている。

【0076】受信テーブルおよび送信テーブルの“packets”は、それぞれ、各インタフェース(interface)から受信したパケット数および各インタフェースへ送信したパケット数を示している。たとえば、インタフェース“eth0”のデータによると、ルータ11Sは、送信側通信ネットワークAから324個のパケットを受信し、267個のパケットを送信したことが示されている。

【0077】受信テーブルおよび送信テーブルの“errs”は、それぞれ、各インタフェースからの受信パケットのエラー数および各インタフェースへの送信パケットのエラー数を示している。

【0078】受信テーブルおよび送信テーブルの“drops”は、それぞれ、各インタフェースからの受信パケットの廃棄数および各インタフェースへの送信パケットの廃棄数を示している。

【0079】受信テーブルおよび送信テーブルの“fifo”は、それぞれ、受信処理待ちのパケット数(キュー(FIFO)の長さ)および送信処理待ちのパケット数(キュー(FIFO)の長さ)を示している。

【0080】受信テーブルの“frame”は受信フレーム数を、“compressed”は圧縮パケット数を、“multicast”は受信マルチキャスト・パケット数をそれぞれ示している。

【0081】送信テーブルの“colls”はイーサネット(CSMA/CD方式)における衝突(collision)の発生数を、“carrier”はイーサネットにおけるキャリア検出数を、“compressed”は圧縮パケット数を、それぞれ示している。

【0082】図6に示すテーブルの“interface type”は、通信ネットワーク1のリンクL1～L3の3つのATMインタフェースのタイプを示し、“0eni”は、インタフェース名“eni”のNo.1であり、“1eni”は、インタフェース名“eni”のNo.2であり、“2eni”は、インタフェース名“eni”のNo.3である。

【0083】“ESI/MAC addr”は、MAC(Media Access Control)アドレスを示している。“AAL(Tx, err, Rx, err, drop)”は、ATMアダプテーション・レイヤ(Adaptation Layer)のパケットに関する統計情報を示している。“Tx”は送信パケット数を、左側の“err”は送信エラー数を、“Rx”は受信パケット数を、右側の“err”は受信エラー数を、“drop”は受信廃棄パケット数を、それぞれ示している。

【0084】図5および図6は、ともにリンクに関するデータであるが、OSは、リンク内に設けられた各LSPに関するデータについても収集することもできる。

【0085】このシステム状態は、OSによりリアルタイム(一定時間間隔 Δt ごと)で収集されているので、この収集された情報に基づいて、前述したリンク・データの各フィールドのデータが求められる。

【0086】たとえば、時刻 $(n-1)$ における図5の送信テーブルの“packets”と、時刻 n における送信テーブルの“packets”との差分を求めることにより、前記式(2)の出力パケット数が求められる。また、この差分である出力パケット数をそのリンクの帯域で除算することにより、式(2)のCUTY(n)が求められる。さらに、このCUTY(n)に基づいて、前記式(1)のAve_Utilizationが求められる。

【0087】また、図5の送信テーブルまたは図6の“drop”は、廃棄パケット数の積算値であるので、時刻 n の“drop”の値が前記式(4)のLoss(n)となる。

【0088】＜オペークLSAのフラッディング方法＞
(1)第1のフラッディング方法

オペークLSAは、フラッディング(flooding)を用いて他のルータに送信される。このフラッディングを行うタイミングは、たとえば、一定時間間隔 T_{th} ごとに行うことができる。この一定時間間隔 T_{th} は、システム状態が更新される前記時間 Δt 以上(好ましくは Δt の正の整数倍)であって、通信ネットワーク1がオペークLSAの送信により輻輳しない間隔であり、かつ、TEを効果的に行うことができる間隔に設定される。その具体的な値は、実験、シミュレーション、計算等により求められる。

【0089】図7は、第1のフラッディング方法の処理の流れを示すフローチャートである。この処理は、処理装置30(図2参照)により実行される。

【0090】まず、ルータ11に電源が投入される等によって、ルータ11が起動すると、処理装置30の内部に設けられたタイマ(図示略)が起動され、計時を開始する(ステップS10)。

【0091】続いて、タイマの時間 t がフラッディング時間間隔 T_{th} に達しているかどうか判定される(ステップS12)。達していない場合には(ステップS12でNO)、他のルータ11からLSAを受信しているかどうか判定される(ステップS26)。他のルータか

らLSAを受信していない場合には(ステップS26でNO)、ステップS12に戻り、受信している場合には(ステップS26でYES)、受信したLSAがオバークLSAかどうか判定される(ステップS28)。この判定は、前述したLSAヘッダのリンク状態タイプの値によって行われる。リンク状態タイプが9、10または11の場合には、オバークLSAと判断され、それ以外の場合には、通常のLSA(オバークLSA以外のLSA)と判断される。

【0092】オバークLSAである場合には(ステップS28でYES)、オバークLSA用データベースの更新処理(後に詳述)が行われ(ステップS30)、その後、受信したオバークLSAが、受信したリンク以外のすべてのリンク(出力リンク)に送出(フラッディング)される(ステップS32)。その後、ステップS12に処理が戻る。一方、通常のLSAである場合には(ステップS28でNO)、通常のLSAの受信処理が行われ(ステップS34)、その後、処理はステップS12に戻る。この通常のLSAの受信処理には、リンク状態データベースの更新処理が含まれている。

【0093】ステップ12でタイマの時間tが時間Th以上である場合には(ステップS12でYES)、他のルータからのLSAが受信されているかどうか判定される(ステップS14)。

【0094】LSAが受信されている場合には(ステップS14でYES)、ステップS16からS20およびS36の処理が行われる。これらの処理は、前述したステップS28からS32およびS34とそれぞれ同じである。その後、タイマ時間tは一定時間Thを経過しているの、自ルータのオバークLSAが生成され、このオバークLSAがすべてのリンク(出力リンク)に送出(フラッディング)される(ステップS22)。

【0095】その後、自ルータのオバークLSAによるオバークLSA用データベースの更新処理が行われる(ステップS23)。そして、タイマがゼロにリセットされ、処理はステップS12に戻る。

【0096】一方、ステップS14において、他のルータからオバークLSAが受信されていない場合には、直ちに自ルータのオバークLSAが生成されて、すべてのリンクに送出された後(ステップS22)、自ルータのオバークLSAによるオバークLSA用データベースの更新処理が行われる(ステップS23)。そして、タイマがリセットされ(ステップS24)、処理は、ステップS12に戻る。

【0097】(2)第2のフラッディング方法
フラッディング方法には、他のルータからオバークLSAを受信すると、その受信したオバークLSAの送信とともに、自己のルータのトラフィック情報を収めたオバークLSAを生成し、送信する方法(この方法を「第2のフラッディング方法」という。)もある。

【0098】図8は、第2のフラッディング方法の処理の流れを示すフローチャートである。前述した図7のフローチャートにおける処理と同じ処理には同じ符号を付し、その詳細な説明を省略することとする。この処理も、処理装置30により実行される。

【0099】まず、ルータ11に電源が投入される等により、ルータ11が起動すると、タイマが起動され(ステップS10)、タイマの時間tが閾値時間Th1に達しているかどうか判定される(ステップS50)。この閾値時間Th1は、ルータ11のOSがシステム状態を更新する時間 Δt 以上であって、かつ、第1のフラッディング方法における時間間隔Thよりも小さな時間に設定されている。この時間Th1をこのように設定するのは、この一定時間Th1を経過する前に、他のルータから2以上のオバークLSAを受信した場合であっても、自ルータについての同一内容のオバークLSAを2回以上フラッディングさせないためである。

【0100】タイマの時間tが閾値時間Th1に達している場合には(ステップS50でYES)、他のルータ11からのオバークLSAの受信を待って、自己のルータ11のトラフィック情報を収めたオバークLSAが作成/送信され(ステップS14~S24、S36)、その後、処理はステップS50に戻る。

【0101】一方、タイマの時間tが閾値時間Th1に達していない場合には(ステップS50でNO)、他のルータ11からオバークLSAを受信しても、自己のルータ11のトラフィック情報を収めたオバークLSAは作成/送信されず(ステップS26~S34)、その後、処理はステップS50に戻る。

【0102】(3)第3のフラッディング方法
第1のフラッディング方法と第2のフラッディング方法とを複合させたフラッディング方法もある。すなわち、オバークLSAを一定時間間隔Thで送信する(第1のフラッディング方法)とともに、他のルータ11からオバークLSAを受信すると、その受信したオバークLSAの送信とともに、自己のルータのトラフィック情報を収めたオバークLSAを生成し、送信する(第2のフラッディング方法)方法である。

【0103】図9および図10は、第3のフラッディング方法の処理の流れを示すフローチャートである。前述した図7および図8のフローチャートにおける処理と同じ処理には同じ符号を付し、その詳細な説明を省略することとする。この処理も、処理装置30により実行される。

【0104】タイマのスタート(ステップS10)後、タイマの時間tが、第2のフラッディング方法で述べた一定時間Th1を経過しているかどうか判定される(ステップS50)。

【0105】タイマ時間tが一定時間Th1に達していない場合には(ステップS50でNO)、ステップS2

6からS34までの処理が実行される。すなわち、他のルータ11からオバークルSAが受信されると、このオバークルSAによりデータベース更新処理が実行されるとともに、このオバークルSAのさらに他のルータに向けて送信する。その後、処理は、ステップS50に戻る。

【0106】一方、タイマ時間tが一定時間間隔Tth1に達している場合には（ステップS50でYES）、タイマ時間tが、第1のフラッシング方法で述べた一定時間間隔Tthと比較される（ステップS12）。

【0107】この一定時間間隔Tthに達していない場合には（ステップS12でNO）、他のルータ11からのLSAの受信を待って、この他のルータ11から受信されたLSAの（データベース更新処理、他のルータへの送信処理等）が実行される（ステップS16からS24、S36）。その後、処理は、ステップS50に戻る。

【0108】一方、一定時間間隔Tthに達している場合には（ステップS12でYES）、他のルータからのオバークルSAの受信の有無に関係なく、自己のルータのトラフィックの状態を格納したオバークルSAが作成され、送信される（ステップS52～S64）。なお、ステップS52～S60の処理は、ステップS26～S34の処理とそれぞれ同じであり、ステップS62～S64の処理は、ステップS22～S24の処理と同じである。その後、処理は、ステップS50に戻る。

【0109】＜オバークルSA用データベースのデータ構造およびその更新処理＞図12は、ルータ11の記憶装置32に設けられたオバークルSA用データベース32bのデータ構造を示す。

【0110】オバークルSA用データベース32bは、ハッシュ・テーブル50と、各ルータ（自ルータを含む。）11のオバークルSAを有するオバークルSAデータベース構造体（以下「OLDB構造体」という。）51a～51c、52a～52d、53aおよび53b等とから構成されている。

【0111】ハッシュ・テーブル50は、複数の記憶セルを備え、各記憶セルは、受信されたオバークルSAのLSAヘッダにある「広告ルータ」の値（すなわち広告ルータのIPアドレス）をハッシュしたハッシュ値h1～h3等をアドレスとしてアクセスされる。各記憶セルには、OLDB構造体へのポインタが格納される。たとえば、ハッシュ値h1に対応する記憶セルには、OLDB構造体51aへのポインタが、ハッシュ値h2に対応する記憶セルには、OLDB構造体52aへのポインタが、ハッシュ値h3に対応する記憶セルには、OLDB構造体53aへのポインタが、それぞれ格納されている。図示しない他のハッシュ値に対応する記憶セルにも、図示しない他のOLDB構造体へのポインタが格納されている。

【0112】OLDB構造体は、受信したオバークルSAの内容（図3および図4に示す各フィールド）を含むとともに、隣接する次のOLDB構造体へのポインタ（たとえばOLDB構造体51aは次のOLDB構造体51bへのポインタ）、ルートまでの距離（コスト）の計算に必要なフィールド等を含んでいる。1つのOLDB構造体には、1つのルータ11が対応している。

【0113】ハッシュ・テーブル50のハッシュ値としては、たとえば、LSAヘッダの「広告ルータ」値（IPアドレス）を整数とみなし、この整数を素数251で除算した商（整数値）が用いられる。この場合に、ハッシュ・テーブル50は、251個の記憶セルを有し、通信ネットワーク1内の複数のルータ11は、251個のグループに分類されることとなる。そして、251個のグループに分類された各グループに属する複数のルータのOLDB構造体は、たとえばOLDB構造体51a～52cのように、ポインタにより連結される。

【0114】なお、同じグループに属する複数のルータは、IPアドレスが近いので、地理的にも近い位置に配置されていることが多い。

【0115】このように、ハッシュ・テーブルを用いることにより、複数のルータをグループに分類することができるので、後述するように、特定のルータのOLDB構造体を検索する時に、検索の高速化を図ることができる。

【0116】図11は、図7から図10に示すステップS18、S23、S30、およびS64のオバークルSA用データベースの更新処理の詳細な処理の流れを示すフローチャートである。この処理は、各ルータ11の処理装置30により実行される。

【0117】まず、オバークルSAを格納するOLDB構造体の記憶領域を記憶装置32内に確保可能かどうか判定される（ステップS100）。確保可能である場合には（ステップS100でYES）、記憶領域が確保され（ステップS102）、確保された記憶領域に、受信したオバークルSAが格納される（ステップS104）。

【0118】続いて、オバークルSAを受信したルータのIPアドレスがハッシュされ、このハッシュ値をアドレスとするハッシュ・テーブル50の記憶セルが決定される（ステップS106）。続いて、決定された記憶セルにポインタにより連結された1または2以上のOLDB構造体の中から、受信されたオバークルSAの発信ルータのIPアドレスと一致するものがあるかどうか判定（すなわち検索）される（ステップS108）。この検索では、検索範囲がハッシュ・テーブル50の1つの記憶セルに対応するOLDB構造体のグループに絞られているので、高速に検索を行うことができる。

【0119】一致するOLDB構造体がない場合には（ステップS118でNO）、受信されたオバークルS

Aを格納したOLDB構造体は、ステップS106で決定された記憶セルに連結されたOLDB構造体の最後尾に連結される。たとえば、ハッシュ値が、図12のハッシュ値h1の場合には、受信されたオバークルSAのOLDB構造体は、OLDB構造体51cの後ろに連結され、OLDB構造体51c内に設けられたポインタには、連結された新たなOLDB構造体へのポインタが格納される。

【0120】一方、ステップS108において、一致するものがある場合には（ステップS108でYES）、受信されたオバークルSAが最新（時間的に最も後）のもの（すなわち、最新のトラフィック情報を有するオバークルSA）かどうか判定される（ステップS110）。

【0121】最新でない場合には（ステップS110でNO）、受信されたオバークルSAは廃棄され、ステップS102で確保された記憶領域は解放される（ステップS112）。一方、最新である場合には（ステップS110でYES）、この最新のOLDB構造体が、データベース32b内に既にある古いOLDB構造体と入れ替えられる（ステップS114）。たとえば、図12のOLDB構造体51bが、受信された新たなオバークルSAを有するOLDB構造体（「OLDB構造体x」とする。）によって入れ替えられる場合には、この同じ位置にOLDB構造体xが挿入される。あるいは、OLDB構造体51cをOLDB構造体51aの直後に連結す

$$\rho_{\text{effective_path_i}} = \rho_{\text{path_i}} \times f(\text{Loss_path_i}) \quad \cdots (8)$$

$$\rho_{\text{effective_path_i}} = \text{Min}(\rho_{\text{effective_path_i}}, \rho_{\text{ceiling}}) \quad \cdots (9)$$

ここで、 $\rho_{\text{path_i}}$ は、経路i（path_i）を構成する1または2以上のリンク（link_j）全体の平均使用率

$$\rho_{\text{path_i}} = \text{Average}(\text{Ave_Utilization}(\text{link_j}, \text{path_i})) \quad \cdots (10)$$

また、Loss_path_iは、経路iを構成する各リンクの廃棄パケット数（Loss_link_j）の合計であり、以下の式により求められる。

【0129】

$$\text{Loss_path_i} = \sum \text{Loss_link_j} \quad \cdots (11)$$

関数fは、パケットの廃棄が発生した場合に、負荷を高めに計算するように修正するための関数である。これは、パケットの廃棄がない場合には、リンクの負荷 $\rho_{\text{path_i}}$ と実効負荷とは一致するが、パケットの廃棄が発生すると、負荷を高めに修正する必要があるためである。 ρ_{ceiling} は、実効負荷の上限値である。

【0130】今、経路1、2、3の実効負荷を、それぞれ

$$\rho_{\text{ave_effective}} = \sum_i (\rho_{\text{effective_path_i}} \times \text{LBW_path_i}) / \sum \text{LBW_path_i} \quad \cdots (12)$$

ここで、LBW_path_iは、経路iの論理帯域である。

【0133】式（12）を上記例に当てはめると、 $\rho_{\text{ave_effective}} = (5\text{M} + 1.6\text{M} + 0.6\text{M}) / (10\text{M} + 8\text{M} + 2\text{M}) = 0.36$

るとともに、OLDB構造体51cの直後にOLDB構造体xを連結することもできる。

【0122】その後、古いオバークルSAのOLDB構造体の記憶領域は解放され（ステップS116）、処理は終了する。

【0123】＜負荷分散の一例＞最後に、オバークルSA用データベース32bに格納された各ルータ11のトラフィック情報に基づいて行われる負荷分散の一例について説明する。

【0124】経路1、2、3（図1参照）の各論理帯域（LSPの帯域）[bps]をそれぞれ10M、8M、2Mとする。

【0125】入側ルータ11S（処理装置30）は、TEによる負荷分散制御を行うために、まず経路1、2、3のそれぞれの実効負荷を算出する。ここで、「実効負荷」とは、リンクの使用率とこのリンクにおけるパケット廃棄率（パケット損失率）とから計算される実効的な使用率である。本来的には、リンクの実際の負荷を計測すればよいが、実際には、ルータ11によって、その内部が多段スイッチ構成になっている場合があり、この場合には、直接的に負荷を計測することは困難であることから、この実効負荷が用いられる。

【0126】この実効負荷は、経路iの実効負荷を $\rho_{\text{effective_path_i}}$ とすると、たとえば以下の式（8）および（9）で求めることができる。

【0127】

$$\rho_{\text{effective_path_i}} = \rho_{\text{path_i}} \times f(\text{Loss_path_i}) \quad \cdots (8)$$

$$\rho_{\text{effective_path_i}} = \text{Min}(\rho_{\text{effective_path_i}}, \rho_{\text{ceiling}}) \quad \cdots (9)$$

であり、以下の式により求められる。

【0128】

$$\rho_{\text{path_i}} = \text{Average}(\text{Ave_Utilization}(\text{link_j}, \text{path_i})) \quad \cdots (10)$$

れ $\rho_{\text{effective_path_1}} = 0.5$ 、 $\rho_{\text{effective_path_2}} = 0.2$ 、 $\rho_{\text{effective_path_3}} = 0.3$ とすると、経路1、2、3の実トラフィック[bps]は、それぞれ $10\text{M} \times 0.5 = 5\text{M}$ 、 $8\text{M} \times 0.2 = 1.6\text{M}$ 、 $2\text{M} \times 0.3 = 0.6\text{M}$ となる。

【0131】続いて、ルータ11S（処理装置30）は、負荷調整（Load Adjusting）を行う。まず、経路1、2、3のすべてを仮想的な1つのパイプとみなし、このパイプの平均使用率 $\rho_{\text{ave_effective}}$ が以下の式により求められる。

【0132】

となる。

【0134】次に、経路間で移動する実効帯域 $\Delta \text{EBW_path_i}$ [bps]が以下の式により算出される。

【0135】

$$\Delta \text{EBW_path_i} = (\rho_{\text{ave_effective}} - \rho_{\text{effective_path_i}} \times \text{LBW_path_i})$$

… (13)

これを経路 1, 2, 3 のそれぞれについて算出すると、
以下ようになる。

【0136】経路 1 の移動する実効帯域 $\Delta \text{EBW_PATH1} = (0.36 - 0.5) \times 10 \text{ M} = -1.4 \text{ M}$

経路 2 の移動する実効帯域 $\Delta \text{EBW_PATH2} = (0.36 - 0.2) \times 8 \text{ M} = +1.28 \text{ M}$

経路 3 の移動する実効帯域 $\Delta \text{EBW_PATH3} = (0.36 - 0.3) \times 2 \text{ M} = +0.12 \text{ M}$

なお、各経路の移動する実効帯域の総和は、 $-1.4 \text{ M} + 1.28 \text{ M} + 0.12 \text{ M} = 0$ となる。

【0137】この計算結果に基づいて負荷調整を行うと、経路 1 については、 $5 \text{ M} - 1.4 \text{ M} = 3.6 \text{ M}$ 、経路 2 については、 $1.6 \text{ M} + 1.28 \text{ M} = 2.88 \text{ M}$ 、経路 3 については、 $0.6 \text{ M} + 0.12 \text{ M} = 0.72 \text{ M}$ となり、経路 1, 2, 3 について、 $3.6 \times G_r : 2.88 \times G_r : 0.72 \times G_r$ の比で負荷分散（トラフィックの分配）の変更を行えばよいこととなる。なお、 G_r は、負荷調整係数である。

【0138】（付記 1）通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードに設けられ、前記通信路のトラフィックに関する情報を収集するトラフィック情報収集装置であって、自己のノードに接続された通信路のトラフィックに関する情報を収集するトラフィック情報収集部と、前記通信ネットワーク上の通信プロトコルで使用されるメッセージにより、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報を他のノードに送信するトラフィック情報送信部と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報を受信するトラフィック情報受信部と、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、前記トラフィック情報受信部により受信された前記他のノードに接続された通信路のトラフィックに関する情報とを記憶するトラフィック情報記憶部と、を備えているトラフィック情報収集装置。

【0139】（付記 2）付記 1 において、前記トラフィック情報送信部は、あらかじめ定められた一定時間間隔ごとに前記自己のノードに接続された通信路のトラフィックに関する情報を送信する、トラフィック情報収集装置。

【0140】（付記 3）付記 1 において、前記トラフィック情報送信部は、前記他のノードに接続された通信路のトラフィックに関する情報が該他のノードから送信されてきた時に、前記自己のノードに接続された通信路のトラフィックに関する情報を送信する、トラフィック

情報収集装置。

【0141】（付記 4）付記 1 において、前記トラフィック情報送信部は、あらかじめ定められた一定時間間隔ごとに前記自己のノードに接続された通信路のトラフィックに関する情報を送信するとともに、前記他のノードに接続された通信路のトラフィックに関する情報が該他のノードから送信されてきた時に、前記自己のノードに接続された通信路のトラフィックに関する情報を送信する、トラフィック情報収集装置。

【0142】（付記 5）通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードに設けられ、前記通信路のトラフィック情報を収集するトラフィック情報収集装置であって、自己のノードに接続された通信路のトラフィックに関する情報を収集するトラフィック情報収集部と、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報とを記憶するトラフィック情報記憶部と、を備えているトラフィック情報収集装置。

【0143】（付記 6）付記 1 から 5 のいずれか 1 つにおいて、前記他のノードから送信されてきたトラフィックに関する情報を、送信されてきた通信路を除く他の通信路に転送するトラフィック情報転送部をさらに備えている、トラフィック情報収集装置。

【0144】（付記 7）付記 1 から 6 のいずれか 1 つにおいて、前記トラフィックに関する情報が、トラフィックを出力する通信路のトラフィックに関する情報である、トラフィック情報収集装置。

【0145】（付記 8）付記 1 から 7 のいずれか 1 つにおいて、前記トラフィック情報収集部が、あらかじめ定められた一定時間間隔ごとに前記自己のノードに接続された通信路のトラフィックに関する情報を収集する、トラフィック情報収集装置。

【0146】（付記 9）付記 1 から 8 のいずれか 1 つにおいて、前記トラフィックに関する情報が、通信路の平均使用率、通信路の廃棄されたバケット数、および通信路の帯域、ならびに、通信路を論理的な通信路に分割した場合における該論理的な通信路の平均使用率、廃棄されたバケット数、および該論理的な通信路の帯域を含む、トラフィック情報収集装置。

【0147】（付記 10）付記 1 から 9 のいずれか 1 つにおいて、前記通信プロトコルで使用されるメッセージが、OSPF におけるオバーク・リンク状態広告である、トラフィック情報収集装置。

【0148】（付記 11）付記 1 から 10 のいずれか 1 つにおいて、前記ノードがルータである、トラフィッ

ク情報収集装置。

【0149】（付記12） 付記1から11のいずれか1つにおいて、前記トラフィック情報収集装置を有するノードが、前記通信ネットワークにおけるトラフィックの負荷分散制御を行うノードである、トラフィック情報収集装置。

【0150】（付記13） 付記1から12のいずれか1つにおいて、前記トラフィック情報記憶部が、ハッシュ・テーブルと、前記各ノードのトラフィックに関する情報を有する構造体データとを備え、前記ハッシュ・テーブルは、前記各ノードをユニークに識別する情報をハッシュしたハッシュ値をアドレスとする記憶セルを有するとともに、各記憶セルには、該記憶セルのハッシュ値に対応するノードのトラフィックに関する情報を有する前記構造体データへのポインタが格納されており、前記構造体データは、同じハッシュ値に対応する他の構造体データがある場合には、該他の構造体データへのポインタをさらに備えている、トラフィック情報収集装置。

【0151】（付記14） 付記1から13のいずれか1つにおいて、前記トラフィック情報記憶部は、トラフィックに関する情報を記憶する際に、同じノードのトラフィックに関する情報が既に存在する場合には、それらのうちの時間的に後のトラフィックに関する情報を記憶する、トラフィック情報収集装置。

【0152】（付記15） 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける該ノードであって、自己に接続された通信路のトラフィックに関する情報を収集するトラフィック情報収集部と、前記通信ネットワーク上の通信プロトコルで使用されるメッセージにより、前記トラフィック情報収集部により収集された前記自己に接続された通信路のトラフィックに関する情報を他のノードに送信するトラフィック情報送信部と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報を受信するトラフィック情報受信部と、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、前記トラフィック情報受信部により受信された前記他のノードに接続された通信路のトラフィックに関する情報とを記憶するトラフィック情報記憶部と、を備えているノード。

【0153】（付記16） 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける該ノードであって、自己に接続された通信路のトラフィックに関する情報を収集するトラフィック情報収集部と、前記トラフィック情報収集部により収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報とを記憶する

トラフィック情報記憶部と、を備えているノード。

【0154】（付記17） 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードで行われる、前記通信路のトラフィックに関する情報を収集するトラフィック情報収集方法であって、自己のノードに接続された通信路のトラフィックに関する情報を収集し、前記通信ネットワーク上の通信プロトコルで使用されるメッセージにより、前記収集された前記自己のノードに接続された通信路のトラフィックに関する情報を他のノードに送信し、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報を受信し、前記収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、前記受信された前記他のノードに接続された通信路のトラフィックに関する情報とを記憶する、トラフィック情報収集方法。

【0155】（付記18） 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードで行われる、前記通信路のトラフィック情報を収集するトラフィック情報収集方法であって、自己のノードに接続された通信路のトラフィックに関する情報を収集し、前記収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報とを記憶する、トラフィック情報収集方法。

【0156】（付記19） 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードで実行される、前記通信路のトラフィックに関する情報を収集するトラフィック情報収集プログラムであって、自己のノードに接続された通信路のトラフィックに関する情報を収集する手順と、前記通信ネットワーク上の通信プロトコルで使用されるメッセージにより、前記収集された前記自己のノードに接続された通信路のトラフィックに関する情報を他のノードに送信する手順と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報を受信する手順と、前記収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、前記受信された前記他のノードに接続された通信路のトラフィックに関する情報とを記憶する手順と、を備えているトラフィック情報収集プログラム。

【0157】（付記20） 通信される情報を送信、受信または中継する複数のノードと、該複数のノードを接続する通信路とを有する通信ネットワークにおける前記ノードで実行される、前記通信路のトラフィック情報を収集するトラフィック情報収集プログラムであって、自

己のノードに接続された通信路のトラフィックに関する情報を収集する手順と、前記収集された前記自己のノードに接続された通信路のトラフィックに関する情報と、他のノードから送信されてきた該他のノードに接続された通信路のトラフィックに関する情報とを記憶する手順と、を備えているトラフィック情報収集プログラム。

【0158】

【発明の効果】本発明によると、通信ネットワーク上のノードが、通信ネットワークにおけるトラフィックに関する情報を収集することができ、該情報により、負荷分散制御を行うことができる。

【図面の簡単な説明】

【図1】本発明に係る情報収集装置を備えたノードを有する通信ネットワーク1の概略構成を示すブロック図である。

【図2】入側ルータの構成を示すブロック図である。

【図3】オペークLSAのデータ構造を示す。

【図4】オペークLSAに含まれるリンク・データのデータ構造を示す。

【図5】ルータのOSがリアルタイムで収集/管理しているシステム状態のデータの一例を示す。

【図6】ルータのOSがリアルタイムで収集/管理しているデータの一例を示す。

【図7】第1のフラッディング方法の処理の流れを示すフローチャートである。

【図8】第2のフラッディング方法の処理の流れを示すフローチャートである。

【図9】第3のフラッディング方法の処理の流れを示すフローチャートである。

【図10】第3のフラッディング方法の処理の流れを示すフローチャートである。

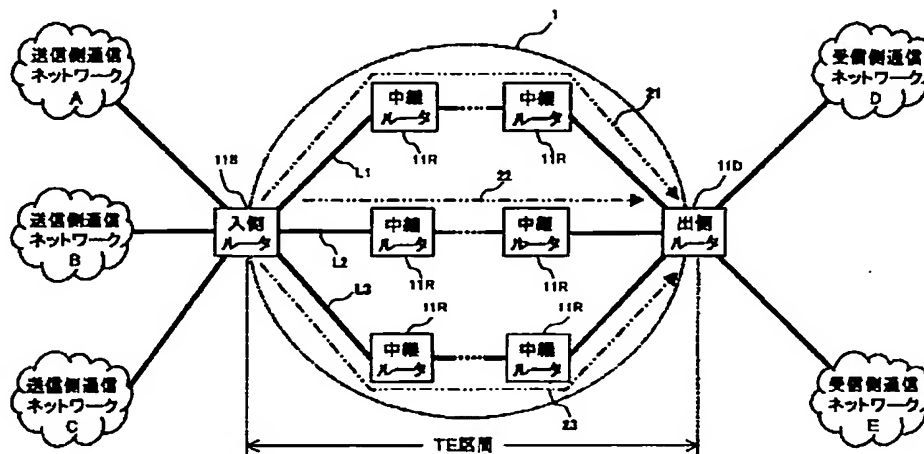
【図11】オペークLSA用データベース更新処理の流れを示すフローチャートである。

【図12】オペークLSA用データベースのデータ構造を示す。

【符号の説明】

- 1 通信ネットワーク
- 11S 入側ルータ
- 11R 中継ルータ
- 11D 出側ルータ
- L1, L2, L3 リンク
- 21, 22, 23 経路(仮想コネクション)
- 30 処理装置
- 32 記憶装置
- 32b オペークLSA用データベース
- 41~47 インタフェース装置

【図1】

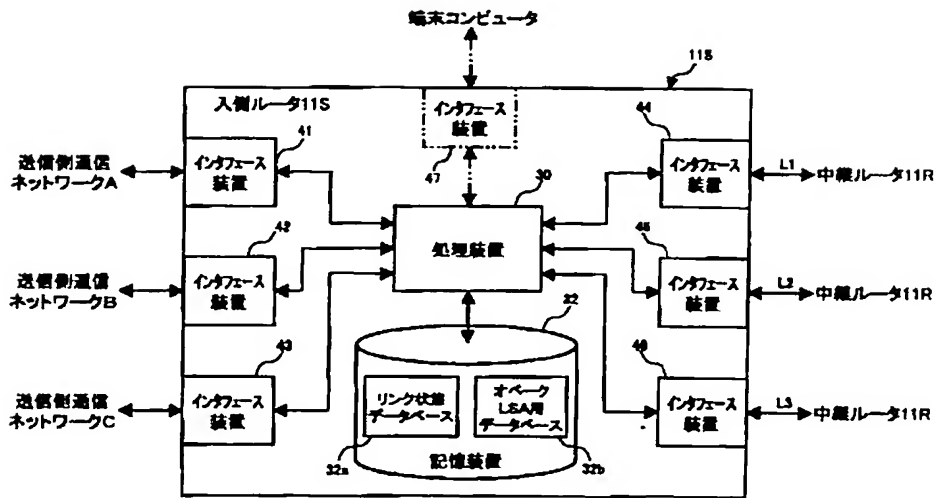


【図6】

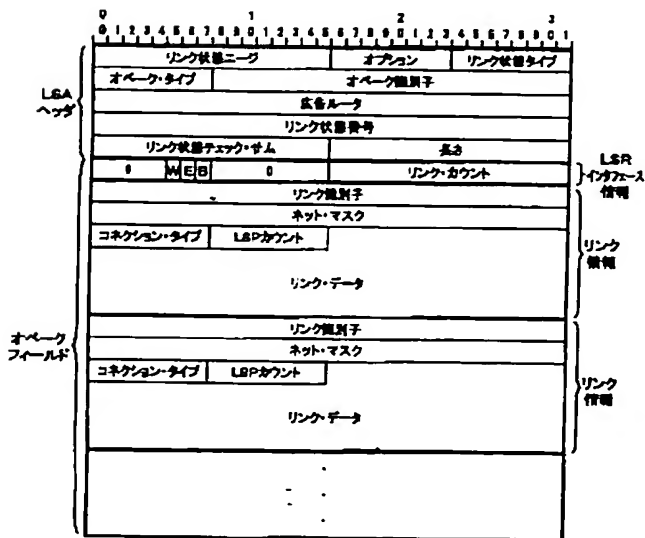
/proc/stm/deviceファイル

interface type	ESI/*MAC* addr	ACL(Tx, err, Rx, err, drop)
0 eni	0020ea00co84	0(0 0 0 0) 0(83 0 76 0 0)
1 eni	0020ea005ee2	0(0 0 0 0) 0(0 0 0 0 0)
2 eni	0020ea005e34	0(0 0 0 0) 0(0 0 0 0 0)

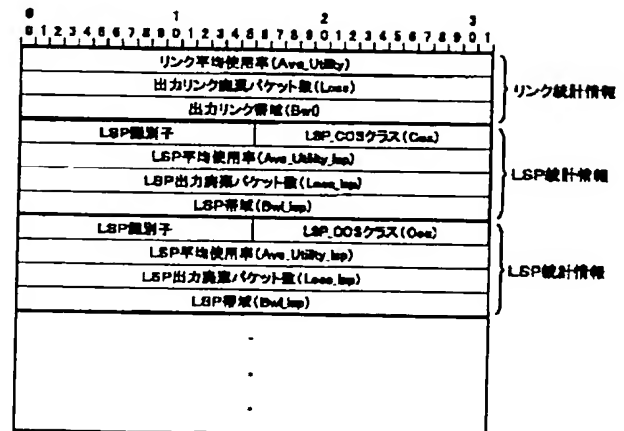
【図2】



【図3】



【図4】

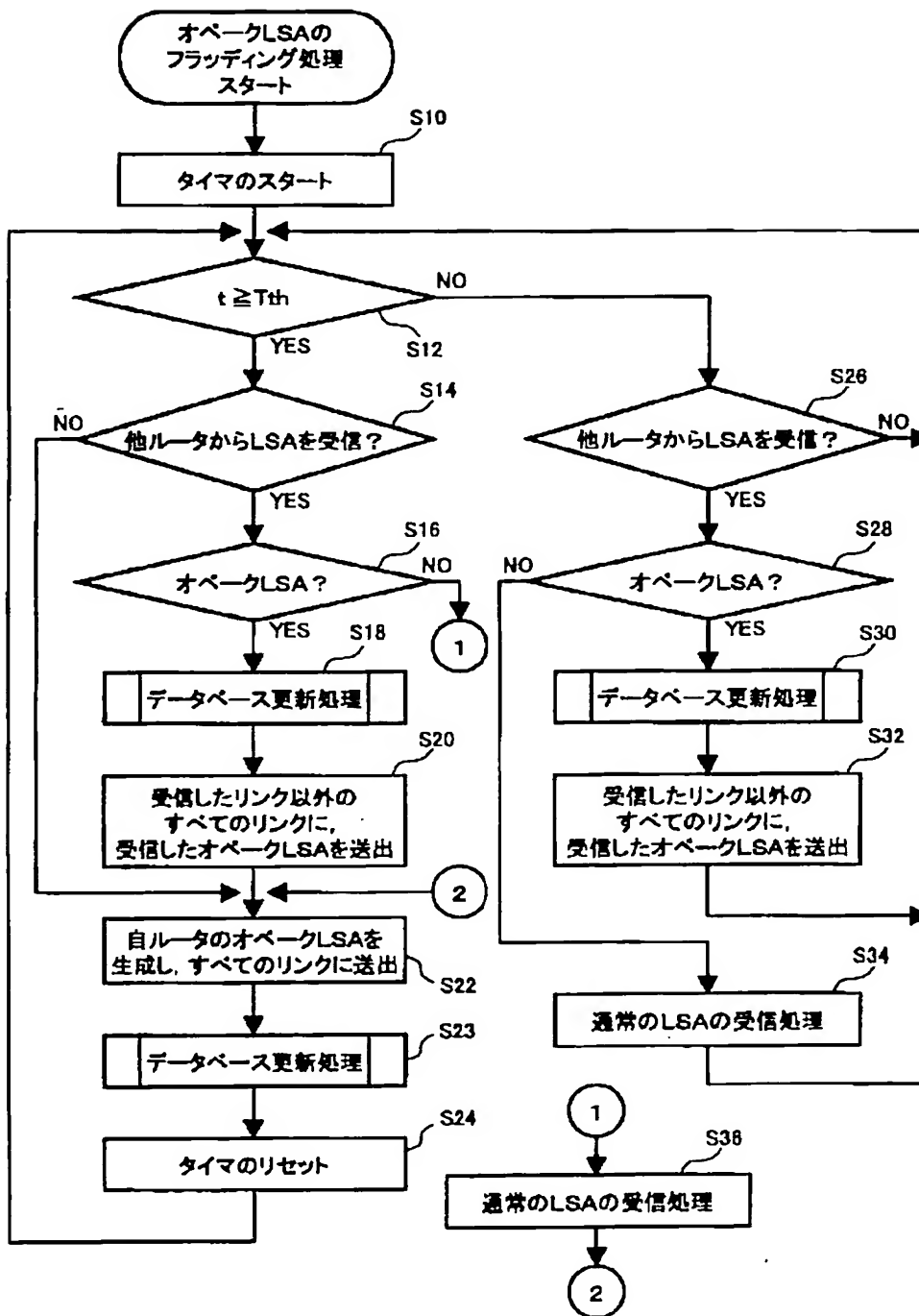


【図 5】

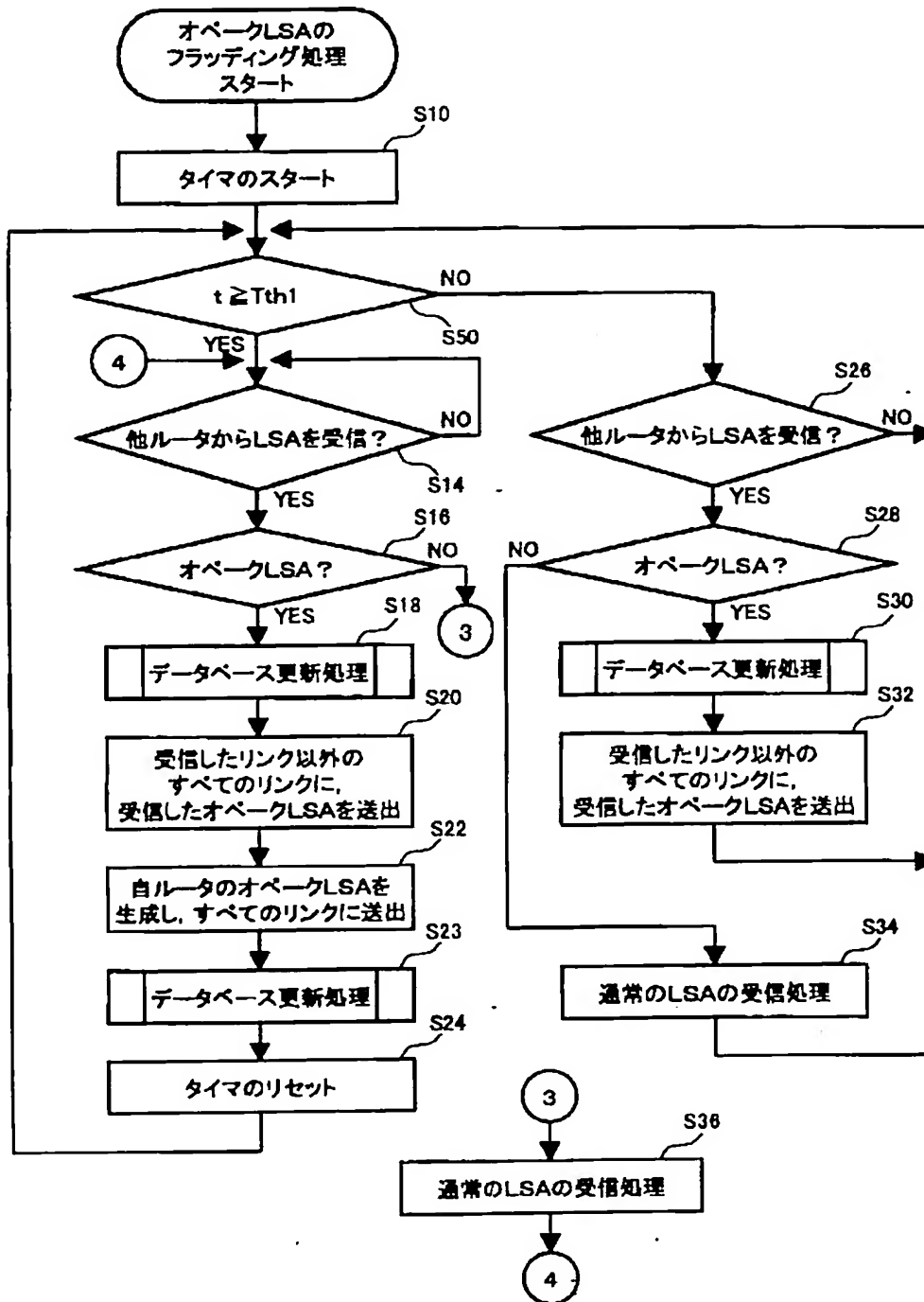
/proc/net/devファイル

interface	Receive							
	bytes	packets	errs	drop	fifo	frame	compressed	multicast
lo	6084	67	0	0	0	0	0	0
eth0	19954	324	0	0	0	0	0	0
eth1	0	0	0	0	0	0	0	0
eth2	0	0	0	0	0	0	0	0
atm0	10488	76	0	0	0	0	0	0
atm1	0	0	0	0	0	0	0	0
atm2	0	0	0	0	0	0	0	0
interface	Transmit							
	bytes	packets	errs	drop	fifo	colls	carrier	compressed
lo	6084	67	0	0	0	0	0	0
eth0	26630	267	0	0	0	0	0	0
eth1	0	0	0	0	0	0	0	0
eth2	0	0	0	0	0	0	0	0
atm0	5054	83	0	0	0	0	0	0
atm1	0	0	0	0	0	0	0	0
atm2	0	0	0	0	0	0	0	0

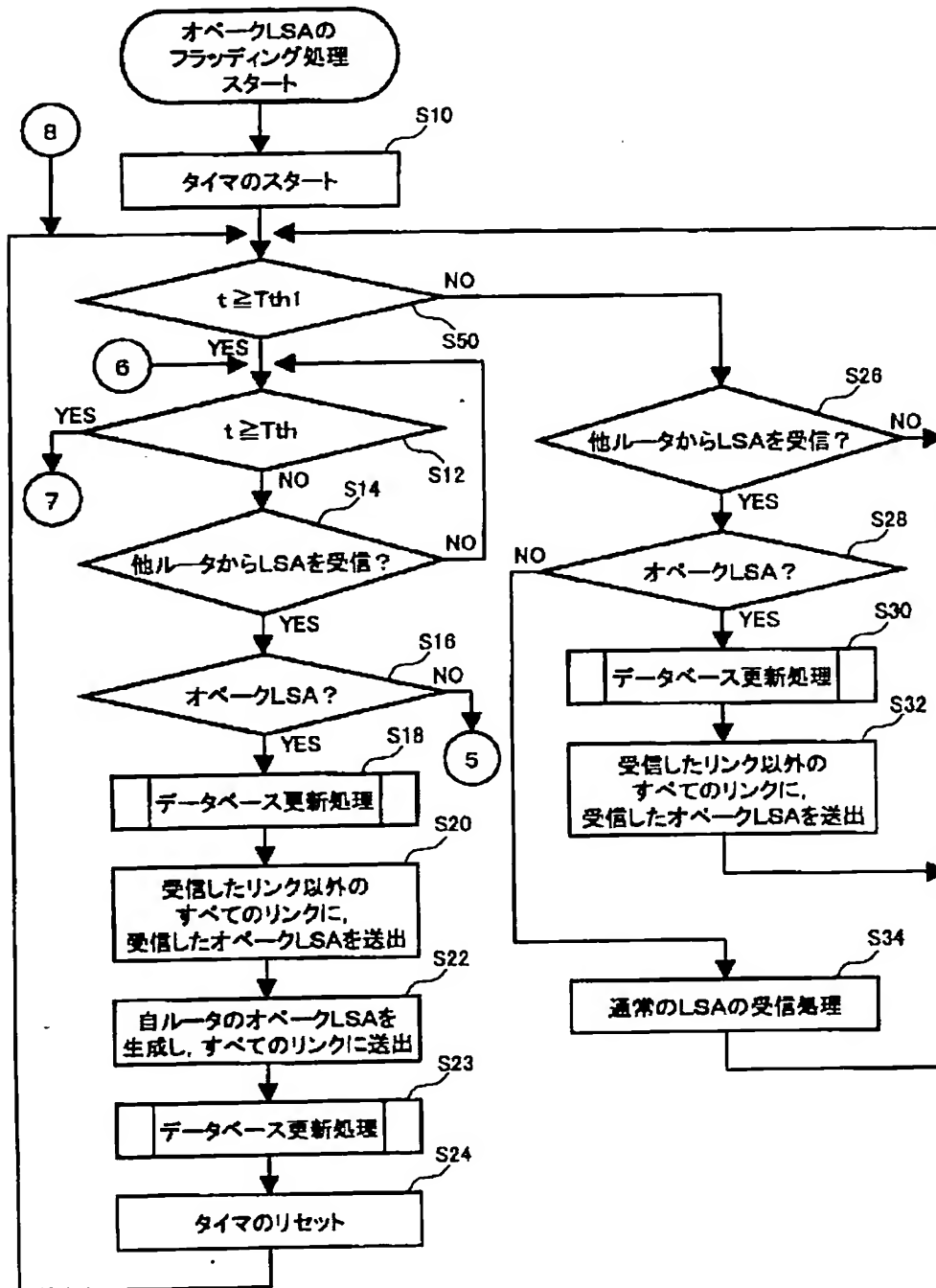
【図 7】



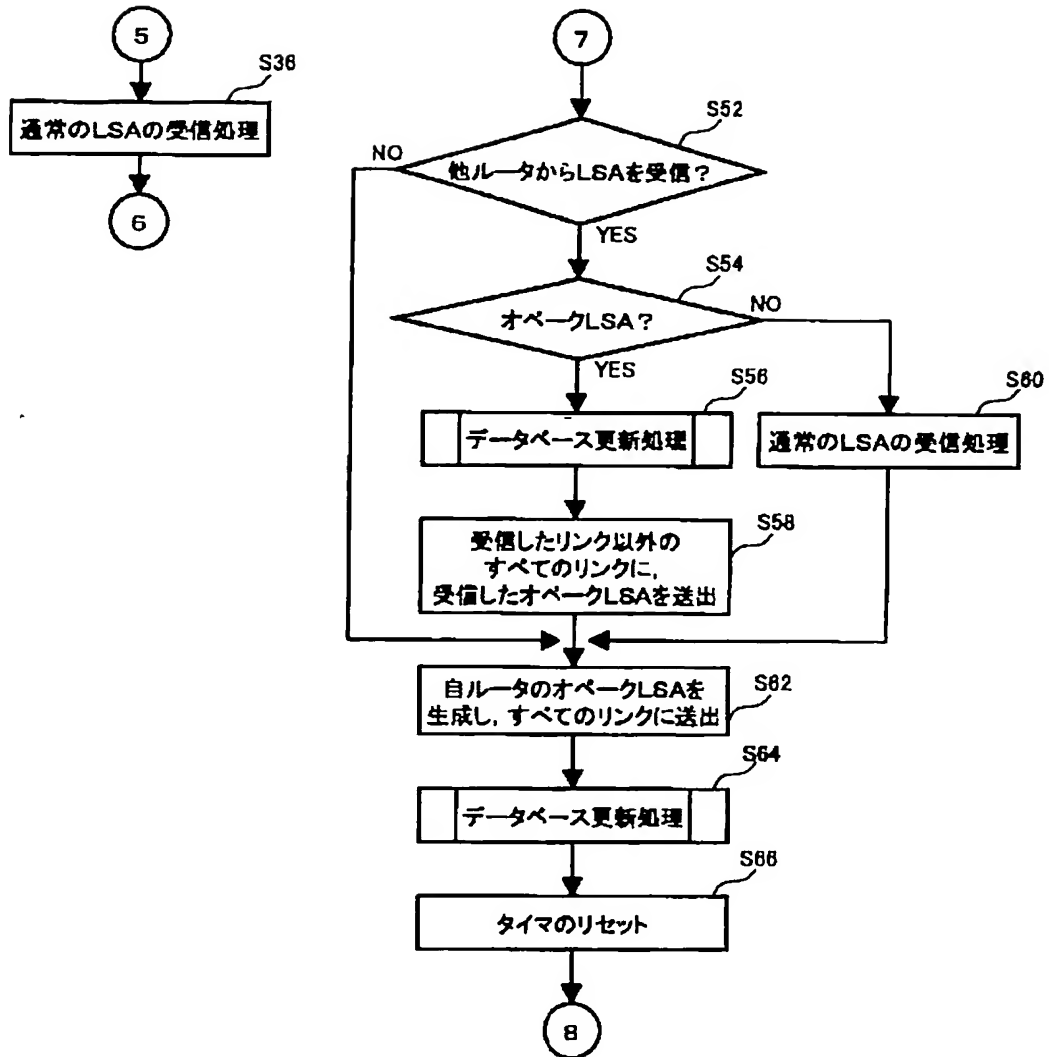
【図 8】



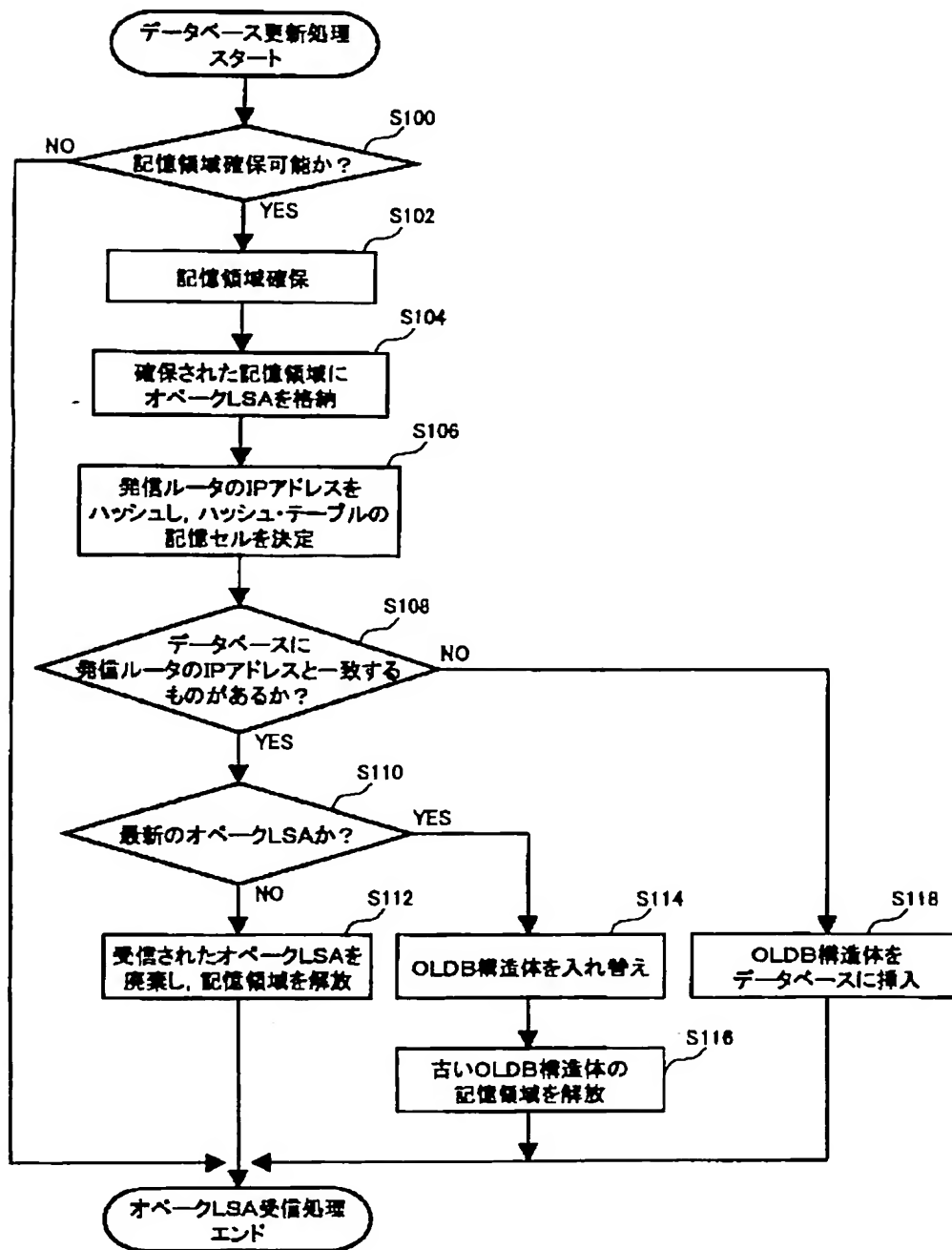
【図 9】



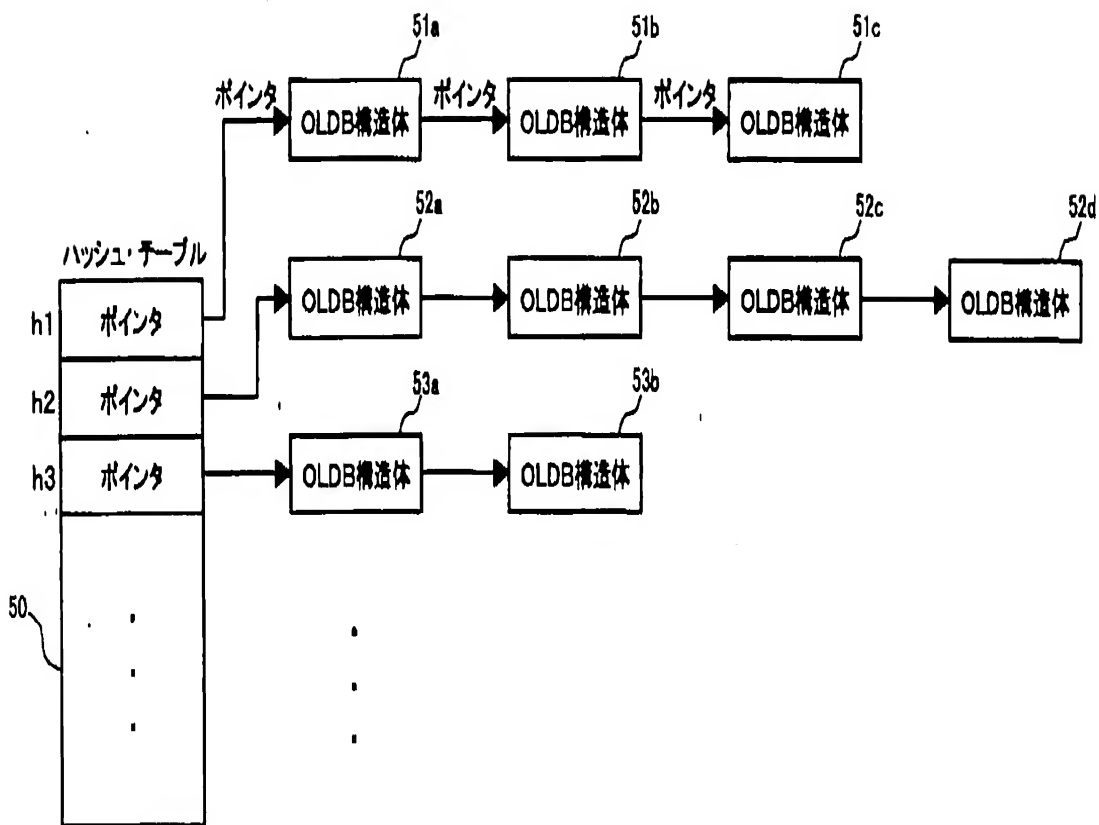
【図 10】



【図 11】



【 図 1 2 】



フロントページの続き

(72) 発明者 宗宮 利夫
神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

Fターム(参考) 5K030 GA13 HA08 HD03 LB05 LE03
MA04 MB09